Writeup Template

Vehicle Detection Project

The goals / steps of this project are the following:

- Perform a Histogram of Oriented Gradients (HOG) feature extraction on a labeled training set of images and train a classifier Linear SVM classifier
- Optionally, you can also apply a color transform and append binned color features, as well as histograms of color, to your HOG feature vector.
- Note: for those first two steps don't forget to normalize your features and randomize a selection for training and testing.
- Implement a sliding-window technique and use your trained classifier to search for vehicles in images.
- Run your pipeline on a video stream (start with the test_video.mp4 and later implement on full project_video.mp4) and create a heat map of recurring detections frame by frame to reject outliers and follow detected vehicles.
- Estimate a bounding box for vehicles detected.

In [1]:

```
# Necessary imports
import matplotlib.image as mpimg
import numpy as np
import cv2
from skimage.feature import hog
import matplotlib.image as mpimg
import matplotlib.pyplot as plt
import numpy as np
import cv2
import glob
import time
from sklearn.svm import LinearSVC
from sklearn.preprocessing import StandardScaler
from skimage.feature import hog
from sklearn.cross validation import train test split
C:\Users\shaurya.dwivedi\AppData\Local\Continuum\Miniconda3\envs\carnd-term
1\lib\site-packages\sklearn\cross validation.py:41: DeprecationWarning: Thi
s module was deprecated in version 0.18 in favor of the model selection mod
ule into which all the refactored classes and functions are moved. Also not
e that the interface of the new CV iterators are different from that of thi
s module. This module will be removed in 0.20.
  "This module will be removed in 0.20.", DeprecationWarning)
```

Histogram of Oriented Gradients (HOG)

- 1. Explain how (and identify where in your code) you extracted HOG features from the training images.
 - I started by reading in all the vehicle and non-vehicle images from the data set provided by udacity.

- I've used all the images from vehicles and non-vehicles.
- I then explored different color spaces and different <code>skimage.hog()</code> parameters (orientations, <code>pixels_per_cell</code>, and <code>cells_per_block</code>). I grabbed random images from each of the two classes and displayed them to get a feel for what the <code>skimage.hog()</code> output looks like.
- I tried different orientations from 8-11 and pixels per cell as 8,16 and 32 while trying different orientations the results were fluctuating a lot.

2. Explain how you settled on your final choice of HOG parameters.

- I tried various combinations of parameters and finally settled with orintation as 9, pixels per cell as 8 and cells per block as 2 which gave the accuracy of my linear svc as 98%.
- While trying different combinations, the test accuracy of my model was fluctuating but after using above combination of parameters i got the accuracy of more than 98% which is pretty decent.

3. Describe how (and identify where in your code) you trained a classifier using your selected HOG features (and color features if you used them).

- I trained a linear SVM using the combination of spatial bin, color histogram and HOG with bin size as 16 and spatial size as (32,32).
- I created the vertical feature stack for the cars and not car images.
- Then i created labels as 1 for the cars and 0 for the non cars.
- I splitted the data as 80-20 for training and tesing and scaled the per column values of training data to avoid overfitting.
- Finally i trained linear SVC as it's pretty fast to train, gives decent accuracy and does the job as well.

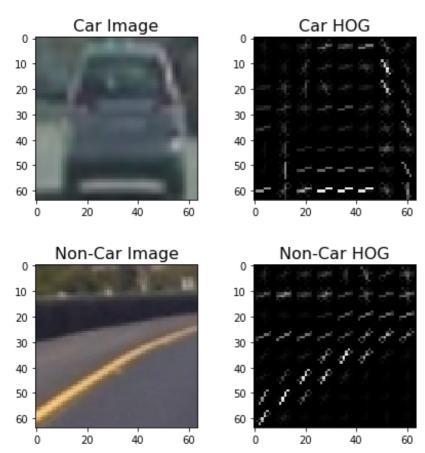
In [2]:

```
# Convert color of the image
def convert color(img, conv='RGB2YCrCb'):
    if conv == 'RGB2YCrCb':
       return cv2.cvtColor(img, cv2.COLOR RGB2YCrCb)
    if conv == 'BGR2YCrCb':
        return cv2.cvtColor(img, cv2.COLOR BGR2YCrCb)
    if conv == 'RGB2LUV':
        return cv2.cvtColor(img, cv2.COLOR RGB2LUV)
# Generate hog features
def get hog features (img, orient, pix per cell, cell per block,
                        vis=False, feature vec=True):
    # Call with two outputs if vis==True
    if vis == True:
        features, hog image = hog(img, orientations=orient,
                                  pixels per cell=(pix per cell,
pix per cell),
                                  cells per block=(cell per block,
cell per block),
                                  block norm= 'L2-Hys',
                                  transform sqrt=False,
                                  visualise=vis, feature vector=feature vec)
        return features, hog_image
    # Otherwise call with one output
    else:
```

```
features = hog(img, orientations=orient,
                       pixels per cell=(pix per cell, pix per cell),
                       cells per block=(cell per block, cell per block),
                       block norm= 'L2-Hys',
                       transform sqrt=False,
                       visualise=vis, feature vector=feature vec)
        return features
# Generate spatial bin
def bin spatial(img, size=(32, 32)):
    color1 = cv2.resize(img[:,:,0], size).ravel()
    color2 = cv2.resize(img[:,:,1], size).ravel()
    color3 = cv2.resize(img[:,:,2], size).ravel()
    return np.hstack((color1, color2, color3))
# Generate color histogram
def color hist(img, nbins=32):
    # Compute the histogram of the color channels separately
    channel1 hist = np.histogram(img[:,:,0], bins=nbins)
    channel2 hist = np.histogram(img[:,:,1], bins=nbins)
    channel3 hist = np.histogram(img[:,:,2], bins=nbins)
    # Concatenate the histograms into a single feature vector
    hist features = np.concatenate((channel1 hist[0], channel2 hist[0], cha
nnel3 hist[0]))
    # Return the individual histograms, bin centers and feature vector
    return hist features
In [3]:
# Read in cars and notcars
vehicles = glob.glob('../vehicles/vehicles/*/*.png')
nonvehicles = glob.glob('../non-vehicles/non-vehicles/*/*.png')
print(len(vehicles))
cars = []
notcars = []
for image in nonvehicles:
   notcars.append(image)
for image1 in vehicles:
    cars.append(image1)
print(len(cars))
8792
8792
8792
In [26]:
car img = mpimg.imread(cars[5])
, car dst = get hog features(car img[:,:,2], 9, 8, 8, vis=True, feature vec
noncar img = mpimg.imread(notcars[5])
_, noncar_dst = get_hog_features(noncar_img[:,:,2], 9, 8, 8, vis=True, featu
re vec=True)
# Visualize
f_{1}((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize=(7,7))
f.subplots adjust(hspace = .4, wspace=.2)
and imphantant im
```

```
ax1.1msnow(car_img)
ax1.set_title('Car Image', fontsize=16)
ax2.imshow(car_dst, cmap='gray')
ax2.set_title('Car HOG', fontsize=16)
ax3.imshow(noncar_img)
ax3.set_title('Non-Car Image', fontsize=16)
ax4.imshow(noncar_dst, cmap='gray')
ax4.set_title('Non-Car HOG', fontsize=16)
print('...')
```

. . .

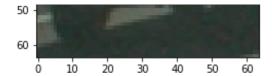


In [23]:

```
# Selecting random car
car = np.random.randint(0, len(cars))
ncar = np.random.randint(0, len(notcars))

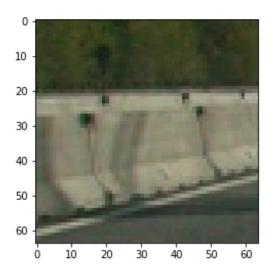
# Display
car_image = mpimg.imread(cars[car])
ncar_image = mpimg.imread(notcars[ncar])
plt.imshow(car_image)
plt.show()
plt.imshow(ncar_image)
```





Out[23]:

<matplotlib.image.AxesImage at 0x1d7c9a47ef0>



In [5]:

```
### TODO: Tweak these parameters and see how the results change.
color space = 'YCrCb' # Can be RGB, HSV, LUV, HLS, YUV, YCrCb
orient = 9 # HOG orientations
pix per cell = 8 # HOG pixels per cell
cell per block = 2 # HOG cells per block
hog channel = "ALL" # Can be 0, 1, 2, or "ALL"
spatial_size = (16, 16) # Spatial binning dimensions
hist bins = 16
               # Number of histogram bins
spatial feat = True # Spatial features on or off
hist feat = True # Histogram features on or off
hog feat = True # HOG features on or off
y start stop = [400, 650] # Min and max in y to search in slide window()
ystart = 400
ystop = 650
scale = 1.5
```

Feature Extraction

Here the feature extraction is done using spatial bin, color histograms and HOG

In [6]:

```
# Read in each one by one
        image = mpimg.imread(file)
        # apply color conversion if other than 'RGB'
        if color space != 'RGB':
            if color space == 'HSV':
                feature image = cv2.cvtColor(image, cv2.COLOR RGB2HSV)
            elif color space == 'LUV':
                feature image = cv2.cvtColor(image, cv2.COLOR RGB2LUV)
            elif color space == 'HLS':
                feature image = cv2.cvtColor(image, cv2.COLOR RGB2HLS)
            elif color space == 'YUV':
                feature image = cv2.cvtColor(image, cv2.COLOR RGB2YUV)
            elif color space == 'YCrCb':
                feature image = cv2.cvtColor(image, cv2.COLOR RGB2YCrCb)
        else: feature image = np.copy(image)
        if spatial feat == True:
            spatial features = bin spatial(feature image, size=spatial size)
            file features.append(spatial features)
        if hist feat == True:
            # Apply color hist()
            hist features = color hist(feature image, nbins=hist bins)
            file features.append(hist features)
        if hog_feat == True:
        # Call get hog features() with vis=False, feature vec=True
            if hog channel == 'ALL':
                hog features = []
                for channel in range(feature image.shape[2]):
                    hog features.append(get hog features(feature image[:,:,
hannel],
                                         orient, pix per cell, cell per bloc
                                         vis=False, feature vec=True))
                hog features = np.ravel(hog features)
            else:
                hog features = get hog features (feature image[:,:,hog channe
1], orient,
                            pix per cell, cell per block, vis=False,
feature vec=True)
            # Append the new feature vector to the features list
            file features.append(hog features)
        features.append(np.concatenate(file features))
    # Return list of feature vectors
    return features
```

In [7]:

```
spatial size=spatial size, hist bins=hist bins,
                        orient=orient, pix per cell=pix per cell,
                        cell per block=cell per block,
                        hog_channel=hog_channel, spatial feat=spatial feat,
                        hist feat=hist feat, hog feat=hog feat)
t2 = time.time()
print(round(t2-t, 2), 'Seconds to extract features...')
print(len(car features))
print(len(notcar features))
4
                                                                          ◎ ▶
137.5 Seconds to extract features...
8792
8968
In [8]:
# Create an array stack of feature vectors
X = np.vstack((car features, notcar features)).astype(np.float64)
# Define the labels vector
y = np.hstack((np.ones(len(car features)), np.zeros(len(notcar features))))
# Split up data into randomized training and test sets
rand state = np.random.randint(0, 100)
X train, X test, y train, y test = train test split(
    X, y, test size=0.2, random state=rand state)
# Fit a per-column scaler
X scaler = StandardScaler().fit(X train)
\# Apply the scaler to X
X train = X scaler.transform(X train)
X test = X scaler.transform(X test)
print('Using:',orient,'orientations',pix per cell,
    'pixels per cell and', cell per block,'cells per block')
print('Feature vector length:', len(X train[0]))
Using: 9 orientations 8 pixels per cell and 2 cells per block
Feature vector length: 6108
In [9]:
# Use a linear SVC
from sklearn.svm import SVC
#svc = SVC(kernel="linear")
svc = LinearSVC()
# Check the training time for the SVC
t=time.time()
svc.fit(X train, y train)
t2 = time.time()
print(round(t2-t, 2), 'Seconds to train SVC...')
# Check the score of the SVC
print('Test Accuracy of SVC = ', round(svc.score(X test, y test), 4))
# Check the prediction time for a single sample
#pred=svc.predict(X test)
#print("Prediction value is ",pred)
t=time.time()
```

```
25.44 Seconds to train SVC...
Test Accuracy of SVC = 0.9837
In [10]:
```

```
from scipy.ndimage.measurements import label
# Add heat to the pixels
def add heat(heatmap, bbox list):
    # Iterate through list of bboxes
    for box in bbox list:
        # Add += 1 for all pixels inside each bbox
        # Assuming each "box" takes the form ((x1, y1), (x2, y2))
        heatmap[box[0][1]:box[1][1], box[0][0]:box[1][0]] += 1
    # Return updated heatmap
    return heatmap# Iterate through list of bboxes
# Apply threshold to reduce false positive
def apply threshold(heatmap, threshold):
    # Zero out pixels below the threshold
    heatmap[heatmap <= threshold] = 0</pre>
    # Return thresholded map
    return heatmap
# Draw labelled boxes
def draw labeled bboxes(img, labels):
    # Iterate through all detected cars
    for car number in range(1, labels[1]+1):
        # Find pixels with each car number label value
        nonzero = (labels[0] == car number).nonzero()
        # Identify x and y values of those pixels
        nonzeroy = np.array(nonzero[0])
        nonzerox = np.array(nonzero[1])
        # Define a bounding box based on min/max x and y
        bbox = ((np.min(nonzerox), np.min(nonzeroy)), (np.max(nonzerox), np.
max(nonzeroy)))
        # Draw the box on the image
        cv2.rectangle(img, bbox[0], bbox[1], (0,0,255), 6)
    # Return the image
    return img
```

Sliding Window Search

- 1. Describe how (and identify where in your code) you implemented a sliding window search. How did you decide what scales to search and how much to overlap windows?
 - I used HOG sub-smapling because it only has to extract hog features once, for each of a small set of predetermined window sizes (defined by a scale argument), and then can be sub-sampled to get all of its overlaying windows.
 - Each window is defined by a scaling factor that impacts the window size. The scale factor can be set on different regions of the image (e.g. small near the horizon, larger in the center).
 - I used 1.5 scale to impact the window size.
- 2. Show some examples of test images to demonstrate how your pipeline is working. What did

you do to optimize the performance of your classifier?

- Ultimately I searched on two scales using YCrCb 3-channel HOG features plus spatially binned color and histograms of color in the feature vector, which provided a nice result.
- Images can be found below as output.

In [24]:

```
# Define a single function that can extract features using hog sub-
sampling and make predictions
def find cars (img, ystart, ystop, scale, svc, X scaler, orient,
pix per cell, cell per block, spatial size, hist bins):
    bboxes=[]
    draw img = np.copy(img)
    img = img.astype(np.float32)/255
    img tosearch = img[ystart:ystop,xstart:xstop,:]
    ctrans tosearch = convert color(img tosearch, conv='RGB2YCrCb')
    if scale != 1:
        imshape = ctrans tosearch.shape
        ctrans tosearch = cv2.resize(ctrans tosearch, (np.int(imshape[1]/sca
le), np.int(imshape[0]/scale)))
    ch1 = ctrans tosearch[:,:,0]
    ch2 = ctrans tosearch[:,:,1]
    ch3 = ctrans tosearch[:,:,2]
    # Define blocks and steps as above
    nxblocks = (ch1.shape[1] // pix per cell) - cell per block + 1
    nyblocks = (ch1.shape[0] // pix per cell) - cell per block + 1
    nfeat per block = orient*cell per block**2
    # 64 was the orginal sampling rate, with 8 cells and 8 pix per cell
    window = 64
    nblocks per window = (window // pix per cell) - cell per block + 1
    cells per step = 2  # Instead of overlap, define how many cells to step
    nxsteps = (nxblocks - nblocks_per_window) // cells_per_step + 1
    nysteps = (nyblocks - nblocks per window) // cells per step + 1
    # Compute individual channel HOG features for the entire image
   hog1 = get hog features(ch1, orient, pix per cell, cell per block,
feature vec=False)
   hog2 = get hog features(ch2, orient, pix per cell, cell per block,
feature vec=False)
    hog3 = get hog features(ch3, orient, pix per cell, cell per block,
feature vec=False)
    for xb in range(nxsteps):
        for yb in range(nysteps):
           ypos = yb*cells per step
            xpos = xb*cells per step
            # Extract HOG for this patch
            hog feat1 = hog1[ypos:ypos+nblocks per window,
xpos:xpos+nblocks per window].ravel()
           hog_feat2 = hog2[ypos:ypos+nblocks per window,
xpos:xpos+nblocks per window].ravel()
           hog feat3 = hog3[ypos:ypos+nblocks per window,
xpos:xpos+nblocks_per_window].ravel()
            hog features = np.hstack((hog feat1, hog feat2, hog feat3))
```

```
xleft = xpos*pix per cell
            ytop = ypos*pix per cell
            # Extract the image patch
            subimg = cv2.resize(ctrans tosearch[ytop:ytop+window, xleft:xlef
t+window], (64,64))
            # Get color features
            spatial features = bin spatial(subimg, size=spatial size)
            hist features = color hist(subimg, nbins=hist bins)
            # Scale features and make a prediction
            test_features = X_scaler.transform(np.hstack((spatial features,
hist features, hog features)).reshape(1, -1))
            #test features = X scaler.transform(np.hstack((shape feat,
hist feat)).reshape(1, -1))
            test_prediction = svc.predict(test features)
            if test prediction == 1:
                xbox left = np.int(xleft*scale)
                ytop draw = np.int(ytop*scale)
                win draw = np.int(window*scale)
                bboxes.append(((xbox_left+xstart, ytop_draw+ystart),(xbox_left)
ft+win draw+xstart,ytop draw+win draw+ystart)))
                cv2.rectangle(draw img, (xbox left+xstart, ytop draw+ystart),
(xbox left+win draw+xstart, ytop draw+win draw+ystart), (0,0,255),6)
    return draw img, bboxes
ystart = 350
ystop = 656
xstart= 450
xstop=1280
scale = 1.5
```

In [21]:

```
# import
from collections import deque
# Variable to store heat to reduce false positive
avg heat=deque(maxlen=20)
# Process pipeline
def pipeline(image):
    draw img1=np.copy(image)
    out img, bboxes = find cars (image, ystart, ystop, scale, svc, X scaler, o
rient, pix per cell, cell per block, spatial size, hist bins)
    out img1,bboxes1 = find cars(image, 390, 480, 1.0, svc, X scaler, orient
, pix per cell, cell per block, spatial size, hist bins)
    bboxes.extend(bboxes1)
    # Add heat to each box in box list
    heat = np.zeros like(image[:,:,0]).astype(np.float)
    heat = add heat(heat,bboxes)
    # Apply threshold to help remove false positives
    avg heat.append(heat)
```

```
heatmap=np.sum(avg_heat,axis=0)

# Visualize the heatmap when displaying
heatmap = np.clip(heatmap, 0,255)
heatmap = apply_threshold(heatmap,40)

# Find final boxes from heatmap using label function
labels = label(heatmap)
draw_img1 = draw_labeled_bboxes(draw_img1, labels)

return draw_img1

**Teturn draw_img1
```

In [17]:

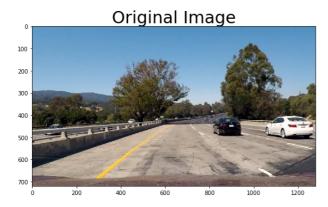
```
#import
from moviepy.editor import VideoFileClip
from IPython.display import HTML
```

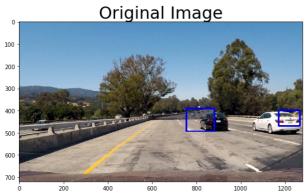
In [22]:

```
#Reading test images
test_images = [mpimg.imread(img) for img in glob.glob("test_images/*.jpg")]
print(len(test_images))

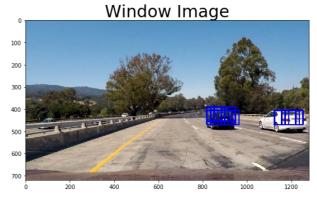
# Plotting test images
for img in test_images:
    plt.imshow(pipeline(img))
    plt.show()
```

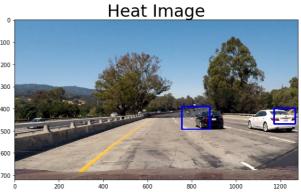
6



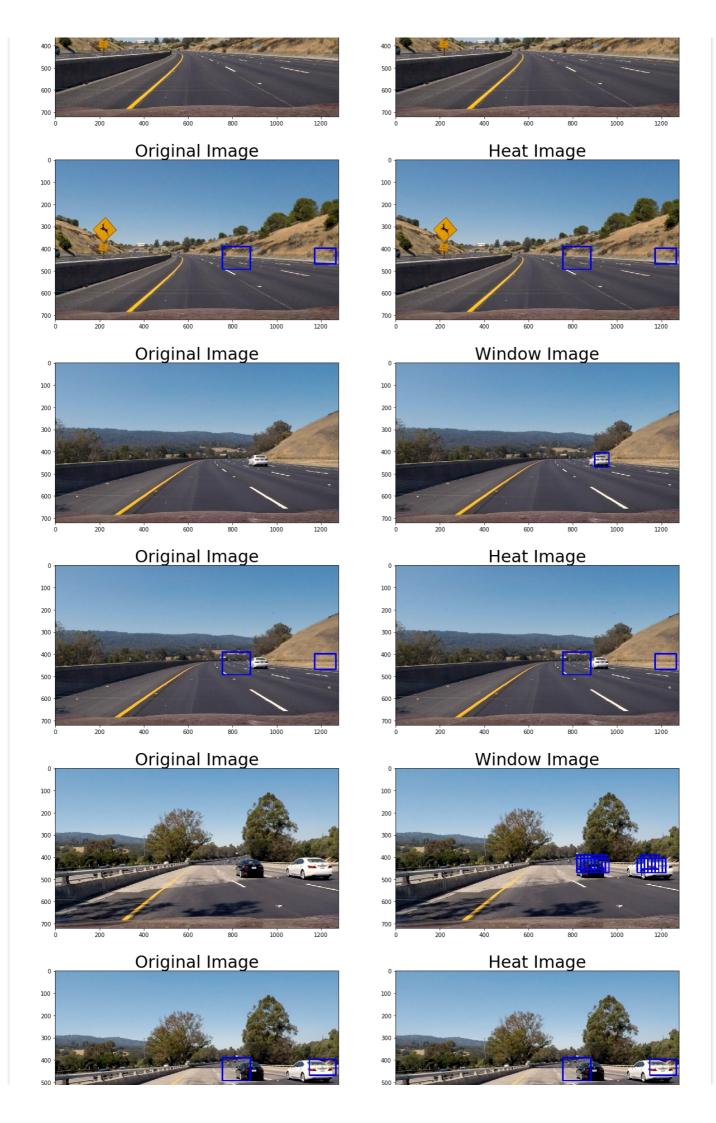


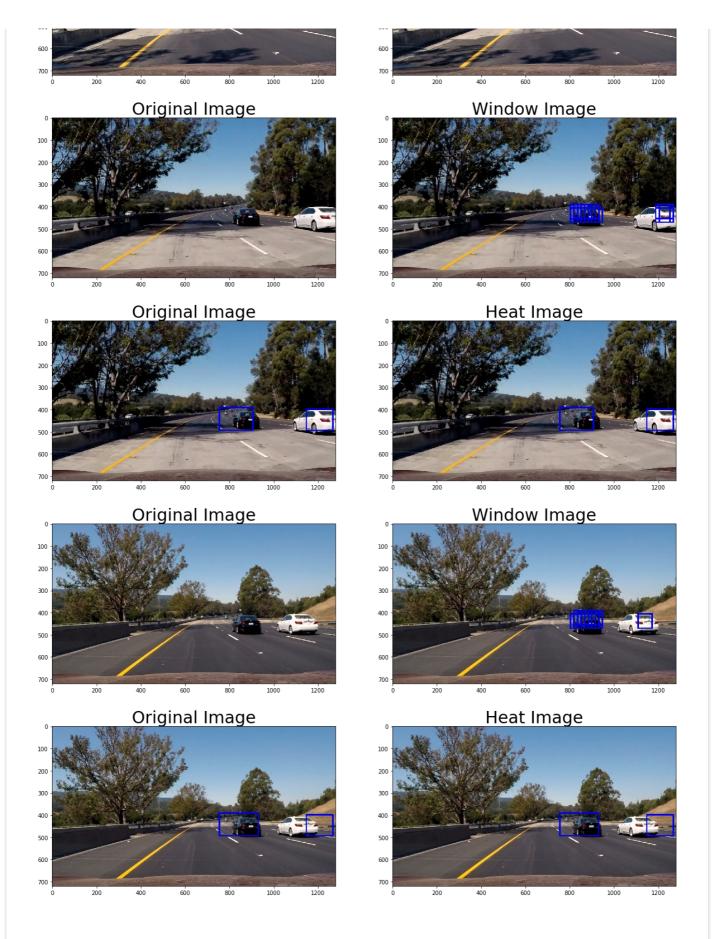












Video Implementation

- 1. Provide a link to your final video output. Your pipeline should perform reasonably well on the entire project video (somewhat wobbly or unstable bounding boxes are ok as long as you are identifying the vehicles most of the time with minimal false positives.)
 - Here's a link to my video result

2. Describe how (and identify where in your code) you implemented some kind of filter for false positives and some method for combining overlapping bounding boxes.

- I recorded the positions of positive detections in each frame of the video. From the positive detections I created a heatmap and then thresholded that map to identify vehicle positions. I then used scipy.ndimage.measurements.label() to identify individual blobs in the heatmap. I then assumed each blob corresponded to a vehicle. I constructed bounding boxes to cover the area of each blob detected.
- The inages can be found above. Due to hardware restrictions i have generated minimum images so that the idea can be plotted.

```
In [27]:
# Output video
white output = 'testvideo output/projectFinal video2.mp4'
clip1 = VideoFileClip("project video.mp4")
white clip = clip1.fl image(pipeline) #NOTE: this function expects color im
ages!!
%time white clip.write videofile(white output, audio=False)
______
                                        Traceback (most recent call last)
<ipython-input-27-ad837308a6dd> in <module>()
      1 # Output video
      2 white output = 'testvideo output/projectFinal video3.mp4'
---> 3 clip1 = VideoFileClip("project video output4.mp4")
      4 white clip = clip1.fl image (pipeline) #NOTE: this function expects
color images!!
      5 get ipython().run line magic('time',
'white clip.write videofile(white output, audio=False)')
~\AppData\Local\Continuum\Miniconda3\envs\carnd-term1\lib\site-packages\mov
iepy\video\io\VideoFileClip.py in __init__(self, filename, has_mask, audio,
audio buffersize, target resolution, resize algorithm, audio fps, audio nby
tes, verbose, fps source)
target resolution=target resolution,
    80
                                               resize algo=resize algoritl
---> 81
                                               fps_source=fps_source)
    82
    83
               # Make some of the reader's attributes accessible from the
clip
~\AppData\Local\Continuum\Miniconda3\envs\carnd-term1\lib\site-packages\mov
iepy\video\io\ffmpeq reader.py in init (self, filename, print infos, buf
size, pix fmt, check duration, target resolution, resize algo, fps source)
     30
               self.filename = filename
     31
               infos = ffmpeg parse infos(filename, print infos, check dura
tion,
---> 32
                                         fps source)
    33
               self.fps = infos['video fps']
               self.size = infos['video size']
~\AppData\Local\Continuum\Miniconda3\envs\carnd-term1\lib\site-packages\mov
iepy\video\io\ffmpeg reader.py in ffmpeg parse infos(filename, print infos,
check duration, fps source)
   254
               popen params["creationflags"] = 0x08000000
   255
```

```
--> 256
            proc = sp.Popen(cmd, **popen params)
    257
    258
            proc.stdout.readline()
~\AppData\Local\Continuum\Miniconda3\envs\carnd-term1\lib\subprocess.py in
 init (self, args, bufsize, executable, stdin, stdout, stderr, preexec fn
, close fds, shell, cwd, env, universal newlines, startupinfo,
creationflags, restore signals, start new session, pass fds)
                         pass fds=()):
                """Create new Popen instance."""
    841
--> 842
                 cleanup()
    843
                # Held while anything is calling waitpid before returncode
has been
    844
                # updated to prevent clobbering returncode if wait() or pol.
() are
~\AppData\Local\Continuum\Miniconda3\envs\carnd-term1\lib\subprocess.py in
cleanup()
    503 def cleanup():
    504
            for inst in active[:]:
--> 505
                res = inst._internal_poll(_deadstate=sys.maxsize)
    506
                if res is not None:
    507
                    trv:
~\AppData\Local\Continuum\Miniconda3\envs\carnd-term1\lib\subprocess.py in
 internal_poll(self, _deadstate, _WaitForSingleObject, _WAIT_OBJECT_0, _Get
ExitCodeProcess)
                    11 11 11
  1257
   1258
                    if self.returncode is None:
-> 1259
                        if WaitForSingleObject(self. handle, 0) == WAIT OF
JECT 0:
   1260
                            self.returncode = GetExitCodeProcess(self. hance)
le)
   1261
                    return self.returncode
OSError: [WinError 6] The handle is invalid
```

Discussion

- 1. Briefly discuss any problems / issues you faced in your implementation of this project. Where will your pipeline likely fail? What could you do to make it more robust?
 - The basic problem was to identify the correct set of parameters that can work. I tried HSV
 color space which gave me almost 100% accuracy but the result while drawing the labelled
 boxes wasn't as desired.
 - The other issue was 'False Positives', to avoid that i used heat map to reduce the unnecessary detections and also summed up the 20 frames for false positive reduction.
 - This pipeline might fail on the images under various lightning conditions and angles. Haven't tried yet but gonna check it's limitation.
 - To make it more robust we can use different color channel combinations as well with different SVM kernels say 'rbf'.
 - The other approach might be to use CNN in place of the classifier. Gonna try that too. :)